

A simple, flexible and high-throughput cloning system for plant genome editing via CRISPR-Cas system

Hyeran Kim[†], Sang-Tae Kim[†], Jahee Ryu¹, Min Kyung Choi¹, Jiyeon Kweon², Beum-Chang Kang¹, Hyo-Min Ahn¹, Suji Bae¹, Jungeun Kim^{1,2}, Jin-Soo Kim^{1,2*} and Sang-Gyu Kim^{1*}

¹Center for Genome Engineering, Institute for Basic Science, Yuseong-gu, Daejeon, 34047, South Korea, ²Department of Chemistry, Seoul National University, Seoul, 08826, South Korea. [†]These authors contributed equally to this work. *Correspondences: Jskim01@snu.ac.kr; sgkim@ibs.re.kr (Kim is fully responsible for distributions of all materials associated with this article)

Abstract CRISPR-Cas9 system is now widely used to edit a target genome in animals and plants. Cas9 protein derived from *Streptococcus pyogenes* (SpCas9) cleaves double-stranded DNA targeted by a chimeric single-guide RNA (sgRNA). For plant genome editing, *Agrobacterium*-mediated T-DNA transformation has been broadly used to express Cas9 proteins and sgRNAs under the control of CaMV 35S and U6/U3 promoter, respectively. We here developed a simple and high-throughput binary vector system to clone a 19–20 bp of sgRNA, which binds to the reverse complement of a target locus, in a large T-DNA binary vector containing an SpCas9 expressing cassette. Two-step cloning procedures: (1) annealing two target-specific oligonucleotides with overhangs specific to the AarI restriction enzyme site of the binary vector; and (2) ligating the annealed oligonucleotides into the two AarI sites of the vector, facilitate the high-throughput production of the positive clones. In addition, Cas9-coding sequence and U6/U3 promoter can be

easily exchanged via the GatewayTM system and unique EcoRI/XhoI sites on the vector, respectively. We examined the mutation ratio and patterns when we transformed these constructs into *Arabidopsis thaliana* and a wild tobacco, *Nicotiana attenuata*. Our vector system will be useful to generate targeted large-scale knock-out lines of model as well as non-model plant.

Keywords: AarI-mediated sgRNA cloning; CRISPR-Cas9 T-DNA binary vector; Exchangeable U6/U3 promoter; Gateway compatible Cas9 cloning

Citation: Kim H, Kim ST, Ryu J, Choi MK, Kweon J, Kang BC, Ahn HM, Bae S, Kim J, Kim JS, Kim SG (2016) A simple, flexible and high-throughput cloning system for plant genome editing via CRISPR-Cas system. *J Integr Plant Biol* XX:XX–XX doi: 10.1111/jipb.12474

Edited by: Jiangqiang Wu, Kunming Institute of Botany, CAS, China

Received Dec. 23, 2015; **Accepted** Mar. 3, 2016

Available online on Mar. 4, 2016 at www.wileyonlinelibrary.com/journal/jipb

© 2016 Institute of Botany, Chinese Academy of Sciences

INTRODUCTION

A microbial immune system using clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) proteins protects prokaryotes from invading viruses by cleaving viral DNA (Deltcheva et al. 2011; Jinek et al. 2012). Type II CRISPR-Cas system from *Streptococcus pyogenes* has been intensively studied and modified to induce targeted mutagenesis in animals and plants (Doudna and Charpentier 2014; Hsu et al. 2014; Kim and Kim 2014; Baltes and Voytas 2015). Engineered *S. pyogenes* Cas9 (SpCas9) endonuclease binds to a single-guide RNA (sgRNA) that mimics the RNA complex of CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA); Cas9-sgRNA complex generates DNA double-strand breaks (DSBs) at the target locus when guided by a 19 or 20 bp sequence of sgRNA (Jinek et al. 2012). DSBs can be then repaired either by homology directed repair (HDR) or non-homologous end-joining (NHEJ) pathway (Rouet et al. 1994; Bibikova et al. 2002). These cellular repair mechanisms can generate precise sequence modification by HDR or introduce variable insertions/deletions (indel) by the error-prone procedure, NHEJ (Kim and Kim 2014; Sternberg and Doudna 2015).

The previous genome editing techniques, zinc finger nuclease (ZFN) and transcription activator-like effector

nuclease (TALEN) rely on sequence-specific DNA binding proteins to guide the FokI nucleases to the targeted genomic locus for generating DSBs (Kim and Kim 2014). Unlike ZFN and TALEN, Cas9 nuclease can recognize the target sites only via an sgRNA. The development of engineered CRISPR-Cas9 system has accelerated targeted genome editing in animals and plants over the last couple of years, because constructing the Cas9 RNA-guided engineered nuclease (RGEN) system is much simpler, faster and cheaper than assembling several ZFs or TALEs for genome editing (Doudna and Charpentier 2014; Hsu et al. 2014; Kim and Kim 2014; Baltes and Voytas 2015).

In addition to the discovery of the native SpCas9 as a standard component of genome editing, there have been developed mutated SpCas9 proteins for several different applications. Cas9 nickase, lacking one of two nuclease active sites via a single substitution (D10A or H840A), has been used to increase the specificity of genome editing (Mali et al. 2013; Ran et al. 2015b). “Dead” Cas9, lacking both nuclease active sites with only target recognition, has been used to perform chromatin immunoprecipitation, targeted gene activation/repression, or more specific genome editing (Qi et al. 2013; Guilinger et al. 2014). To reduce off-target effect, high fidelity Cas9 variants were developed recently by structure-based mutagenesis of SpCas9 (Slaymaker et al. 2016; Kleinstiver et al.

2016). One limitation of SpCas9-mediated genome editing is the requirement for a protospacer adjacent motif (PAM, 5'-NGG-3') adjacent to the 3' end of the target sequence (Mojica et al. 2009). To overcome this limitation, the native SpCas9 protein was modified to recognize different PAM sequences rather than 5'-NGG-3' without any reduction on genome editing efficiency (Kleinstiver et al. 2015). Thus, various repertoires of Cas9 need to be implemented in plant genome editing tools.

Since three seminal reports showed the success of CRISPR-Cas9-mediated plant genome editing (Li et al. 2013; Nekrasov et al. 2013; Shan et al. 2013), this system has been applied to edit major crop genomes (Voytas and Gao 2014; Araki and Ishii 2015; Kim et al. 2015b), such as wheat (Shan et al. 2013; Wang et al. 2014), rice (Shan et al. 2013), potato (Wang et al. 2015), soybean (Sun et al. 2015) and tomato (Brooks et al. 2014). To edit a plant genome, several binary vectors have been constructed to express Cas9 proteins and sgRNAs in plant cells (Belhaj et al. 2013; Xing et al. 2014; Ma et al. 2015). An engineered Cas9 protein contains at least one or two nuclear localization signals (NLS) in its N- or C-terminus and is usually expressed under the CaMV 35S promoter. The promoter of U6 small nuclear RNA (snRNA) for dicots or U3 snRNA for monocots has been widely used to express an sgRNA (Belhaj et al. 2013), because the length of U6/U3 snRNA promoters is relatively short and the termination signal is simple (TTTTT). However, U6/U3 promoter prefers a guanine/adenine to initiate transcription so that the 5' end of sgRNA is fixed with "G" or "A". Interestingly, adding two "G" nucleotides at the 5' end of sgRNA reduces the off-target effect in animal cells (Cho et al. 2014; Kim et al. 2015a).

Known cloning strategies largely follow two steps: (1) the U6 promoter-sgRNA cassette is prepared by overlapping polymerase chain reaction (PCR) or subcloning; (2) this cassette is inserted into Cas9 cassette-containing binary vector (Belhaj et al. 2013; Xing et al. 2014; Ma et al. 2015). Here, we introduce an improved new binary vector system for plant genome editing through CRISPR-Cas9 system and report genome editing results carried out using our vector constructs in *Arabidopsis thaliana* and *Nicotiana attenuata*. Our cloning

steps are relatively simple, fast and effective as all procedures can be finished within a day. Moreover, our cloning system makes it much easier to replace Cas9 cassette via a Gateway™ system (Karimi et al. 2002; Karimi et al. 2007) and to exchange the U6 promoter after double digestion with XhoI and EcoRI in a target plant.

RESULTS

Construction and features of Cas9-sgRNA vectors for plant transformation

To design a new CRISPR-Cas vector for plant transformation, we first considered that a short guide sequence (19 or 20 bp) of an sgRNA can be directly cloned into a binary vector through a single ligation step without any PCR. To do this, we chose the type IIS restriction enzyme AarI, which can cut a 4 and 8 bp outside from its binding sequence 5'-CACCTGC(N)_{4/8}-3' and generate a non-palindromic overhang of any sequence. We added two AarI recognition sites between the U6 promoter and sgRNA scaffold to place a guide sequence precisely after the 3' end of U6 promoter and before the 5' end of sgRNA scaffold (Figure 1). The AarI enzyme allows assembling two DNA overhangs without an extra sequence in the junction. This U6 promoter-AarI-XhoI-AarI-sgRNA scaffold cassette was inserted into Gateway™ binary vectors pB2GW7, harboring the Basta™-resistance gene, and pH2GW7 harboring the hygromycin-resistance gene, developed for *Agrobacterium*-mediated plant transformation (Figure 1) (Karimi et al. 2002). We modified the original pB2GW7 to remove the AarI recognition site in the Basta™-resistance gene. The engineered SpCas9 coding sequence (Cho et al. 2013) was replaced with a standard Gateway™ cassette (attR1-ccdB-attR2) located downstream of the CaMV 35S promoter in both pB(H) 2GW7 vectors (Figure 1).

The *Arabidopsis* U6 or the rice U3 promoter is generally used to express sgRNA in dicot and monocot plants, respectively (Belhaj et al. 2013). However, in some cases U6 promoters derived from the target plants function better than heterologous U6 promoters (Sun et al. 2015). Therefore, we

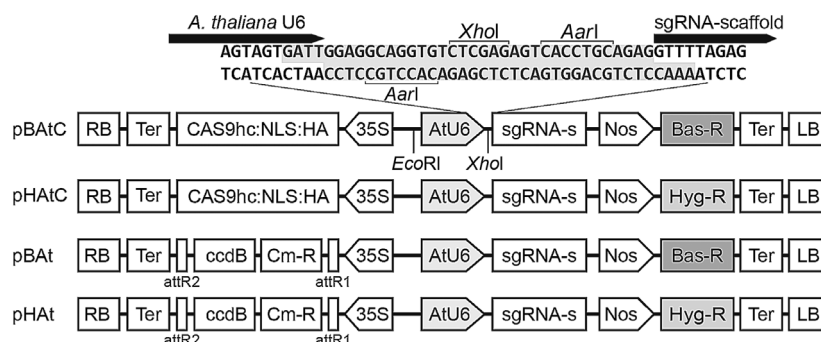


Figure 1. Schematic maps of complete plant RGEN binary vectors

The nomenclature of our vector system is as follows: B, Basta™ resistance gene; H, hygromycin resistance gene; At, *Arabidopsis* U6-26 promoter driven single-guide RNA (sgRNA) cassette; C, human codon-optimized Cas9 expressing cassette. The gray shaded sequence represents nucleotides that are removed after AarI cutting in RGEN binary vector. RB, right border sequences; Ter, terminator sequence; 35S, CaMV 35S promoter; Nos, Nos promoter; LB, left border sequences; RGEN, RNA-guided engineered nuclease.

added the *EcoRI* site at 5' upstream of the U6 promoter and the *XhoI* site between two *AarI* sites at 3' downstream of the U6 promoter to easily replace U6 promoter (Figure 1). The nomenclature of our vector system is as follows: B, Basta™ selection marker; H, Hygromycin selection marker; At, *A. thaliana* U6 promoter; C, engineered SpCas9 (Figures S1, S2).

A simple and fast cloning procedure of an sgRNA

To clone a 20 bp guide sequence (N₂₀) in pB(H)AtC vectors, we designed two single-stranded oligonucleotides: (1) one contains a "GATT" sequence at the 5' end, which is complementary to the 5' overhang generated by the *AarI* digestion of pB(H)AtC vectors at the end of *A. thaliana* U6 promoter, followed by "G" required for transcription initiation by the U6 promoter and a target-specific guide sequence N₂₀; (2) the other oligonucleotide contains a "AAAC" at the 5' end, which is complementary to the 5' overhang generated by the *AarI* digestion at the beginning of sgRNA scaffold, followed by the reverse complement sequence of GN₂₀ (Figure 2A). Annealed products of two oligonucleotides were ligated into *AarI*-cut pB(H)AtC vectors (Figure 2B). The cloned vectors were transformed into *Escherichia coli* competent cells. The cloning efficiency of a

guide sequence was more than 80% of spectinomycin-selected colonies. We normally aliquot *AarI*-cut vectors to a 96-well plate or an eight-well strip tube and use them as "a ready vector" for large-scale cloning with various sgRNAs.

In vivo evaluation of the Cas9-sgRNA vectors for genome editing

To test the functionality of our vectors, we transformed *A. thaliana* and a wild tobacco *N. attenuata* using *Agrobacterium* GV3101 carrying the final vector constructs. We designed candidate sgRNAs using CRISPR RGEN Tools (<http://rgenome.ibs.re.kr>), which provides SpCas9 target sgRNA sequences in the gene of interest with potential off-target sites in any given plant genomes as well as the score predicting microhomology-mediated end joining repair after DSBs (Bae et al. 2014; Park et al. 2015).

We selected three sgRNAs to target the *AtSH3P3* gene (AT4G18060) in *A. thaliana* Col-0, because to date no knockout mutant of *AtSH3P3* has been reported. After selecting hygromycin-resistant T1 plants (Figure 3A), we extracted genomic DNA from young leaf tissues of 10 T1 plants harboring the *AtSH3P3*-sgRNA1 construct and examined the indel ratio and patterns in the T1 plants using the next-generation sequencing (NGS) technology. The indel ratio was calculated

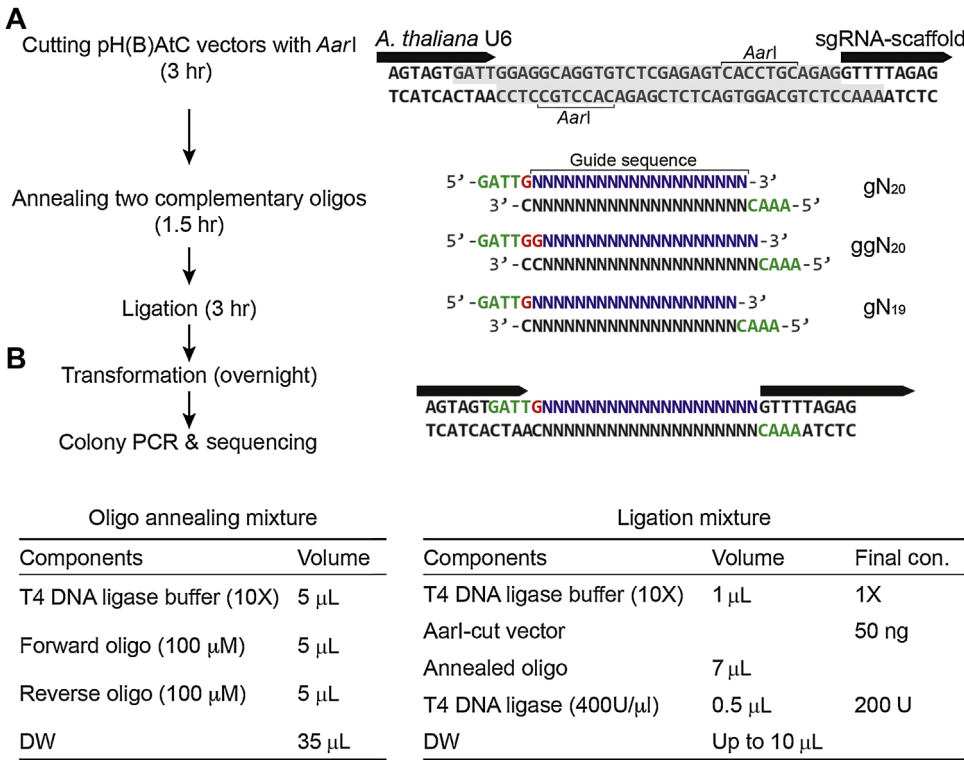


Figure 2. Cloning procedures of a target sgRNA to complete plant RGEN binary vectors

(A) To insert a target binding sequence between *Arabidopsis thaliana* U6 promoter and single-guide RNA (sgRNA) backbone, we design two short oligonucleotides as follows: Sense oligo, 5'-GATTG(N₁₉₋₂₀)-3', where (N₁₉₋₂₀) is the target sequence; antisense oligo, 5'-AAAC(N₁₉₋₂₀)C-3', where (N₁₉₋₂₀) is the complementary sequence to the target sequence. The guanine nucleotide at fifth position (red) in the sense oligo is essential for U6 promoter-driven expression. A pair of annealed oligonucleotides is directly cloned into the *AarI* digested RNA-guided engineered nuclease (RGEN) binary vector: GATT and CAAA nucleotides (green) in annealed oligos are essential for the ligation. The gray box represents nucleotides that are removed after *AarI* cutting in RGEN binary vector. (B) Oligo annealing and ligation mixtures.

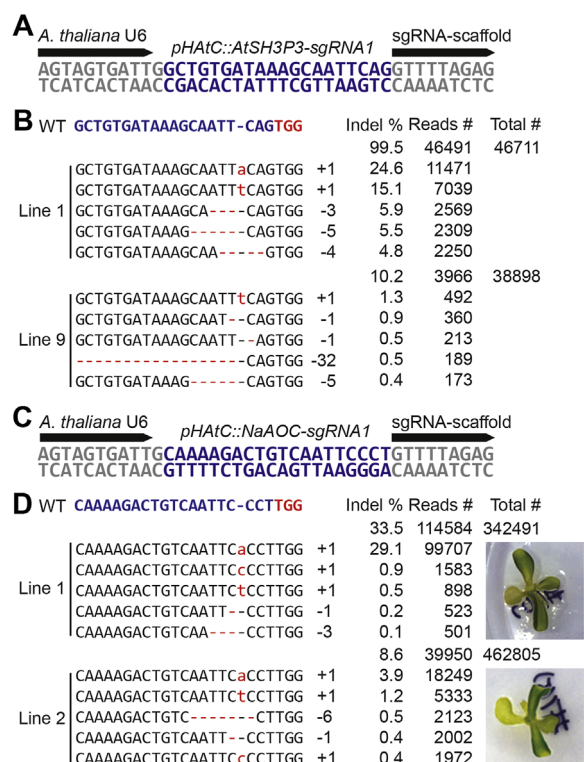


Figure 3. Genome editing in *Arabidopsis thaliana* and *Nicotiana attenuata*

Cloned guide sequences (blue) of single-guide RNA (sgRNA) targeted to (A) *AtSH3P3* gene in *A. thaliana* and (C) *NaAOC* gene in *N. attenuata*. Targeted next-generation sequencing (NGS) analysis of selected (B) T1 *A. thaliana* plants harboring the pHAtC::AtSH3P3-sgRNA1 construct and (D) regenerated *N. attenuata* plantlets harboring the pHAtC::NaAOC-sgRNA1 construct. Wild type (WT) sequences of *AtSH3P3* and *NaAOC* gene are shown in blue with the protospacer adjacent motif sequence (PAM) in red. The indel sequences of target locus are ranked with the read numbers. +; insertion, -; deletion. Indels are marked in red lower case or dash below a WT sequence. The mutation ratio (Indel %) was calculated by dividing the number of sequencing reads containing indel at the target site by the number of total sequencing reads.

by dividing the number of reads harboring indel at *AtSH3P3* target sequence by the number of total reads. The results showed that two lines out of 10 T1 plants had the mutated target gene with various patterns of indel. Line 1 showed a 99.5% indel ratio, and line 9 showed a 10.2% indel ratio (Figure 3B). The major indel pattern of lines 1 and 9 was a single-base insertion at the cleavage site with 39.8% and 1.3% indel ratio, respectively (Figure 3B).

We also transformed a non-model plant, *N. attenuata* using hypocotyl segments inoculated with *Agrobacterium* as described in Krügel et al. (2002). We designed two sgRNAs to edit *N. attenuata* ALLENE OXIDE CYCLASE (*NaAOC*) gene, which is an essential gene for the biosynthesis of jasmonic acid. Transformed hypocotyl tissues were screened on the selection media, and subsequently, genomic DNA was extracted

from regenerated leaves in T0 plants (Figure 3C). We examined the indel ratio and patterns in leaf tissues of two transformed *N. attenuata* plants harboring *NaAOC*-sgRNA1. The total mutation ratio was 33.5% for line 1 and 8.6% for line 2 near the cleavage site (Figure 3D). The major pattern of both lines 1 and 2 was also a single-base insertion at the cleavage site (29.1% for line 1 and 3.9% for line 2; Figure 3D).

To examine whether the mutation ratio observed in somatic tissues in T1 *Arabidopsis* transgenic plants is correlated with that in the next generation, we calculated the mutation ratio in 37 T2 plants derived from two different T1 lines (line 1 and line 9; Figure 3B). Among 19 T2 plants from the line 1 of T1 plant, 11 plants showed higher mutation ratios than that shown in parental T1 plant (99.5%); four plants showed lower than 5% and the other T2 plants showed the range of 36%–88.9% (Figure 4A). Mutation ratios in 11 out of 18 T2 plants from line 2 had higher mutation ratios (13.6%–42.7%) than that in T1 parental plant (10.2%), while other T2 plants showed less than 3% of mutation ratios (Figure 4). Our Sanger sequencing (partially shown in Figure S3) and targeted NGS analysis with T2 plants confirmed that the pHAtC binary vector-mediated genome editing system was successfully established to get inheritable and targeted genome edited plants.

DISCUSSION

We developed a simple and fast AarI-mediated sgRNA cloning system, which facilitates plant genome editing for both targeted and high-throughput gene screening using the CRISPR-Cas9 system. As with our system, another type IIS restriction enzyme, *BsaI* was used to insert one sgRNA into CRISPR-Cas9 vectors (Xing et al. 2014), which were derived from pCambia and pGreen-like vectors. The cloning strategy based on AarI and *BsaI* has neither PCR amplification nor subsequent gel purification steps (except purifying AarI- or *BsaI*-cut vectors), and it can achieve more accurate and efficient cloning for sgRNA as well as high-throughput cloning. In addition to using the AarI type IIS enzyme, we provided GatewayTM platform to easily clone various repertoires of SpCas9. While we used human codon-optimized SpCas9 bearing an NLS and an HA tag at the C-terminus, all kinds of SpCas9 derivatives, such as SpCas9 nickase, “Dead” Cas9, GFP-tagged SpCas9, or engineered SpCas9s, which have different PAM recognition (Kleinstiver et al. 2015) or high fidelity Cas9 (Slaymaker et al. 2016; Kleinstiver et al. 2016), can be easily cloned into our pH(B)AtC binary vectors via GatewayTM compatible exchange, which already harbors the U6-AarI-XhoI-AarI cassette (Figure 1).

Recently, different CRISPR-Cas systems have been intensively examined to broaden the application of this system. For instance, the SaCas9 protein originating from *Staphylococcus aureus* is smaller than SpCas9 and can be transferred to animal cells by adeno-associated virus-mediated delivery (Ran et al. 2015a). In addition, Cpf1 proteins in the putative type V CRISPR-Cas systems generate “sticky ends” at the cleavage sites, which might enhance the HDR-mediated repair system (Zetsche et al. 2015). This cloning system is obviously convenient to clone sgRNAs for SaCas9- or Cpf1-mediated plant genome editing.

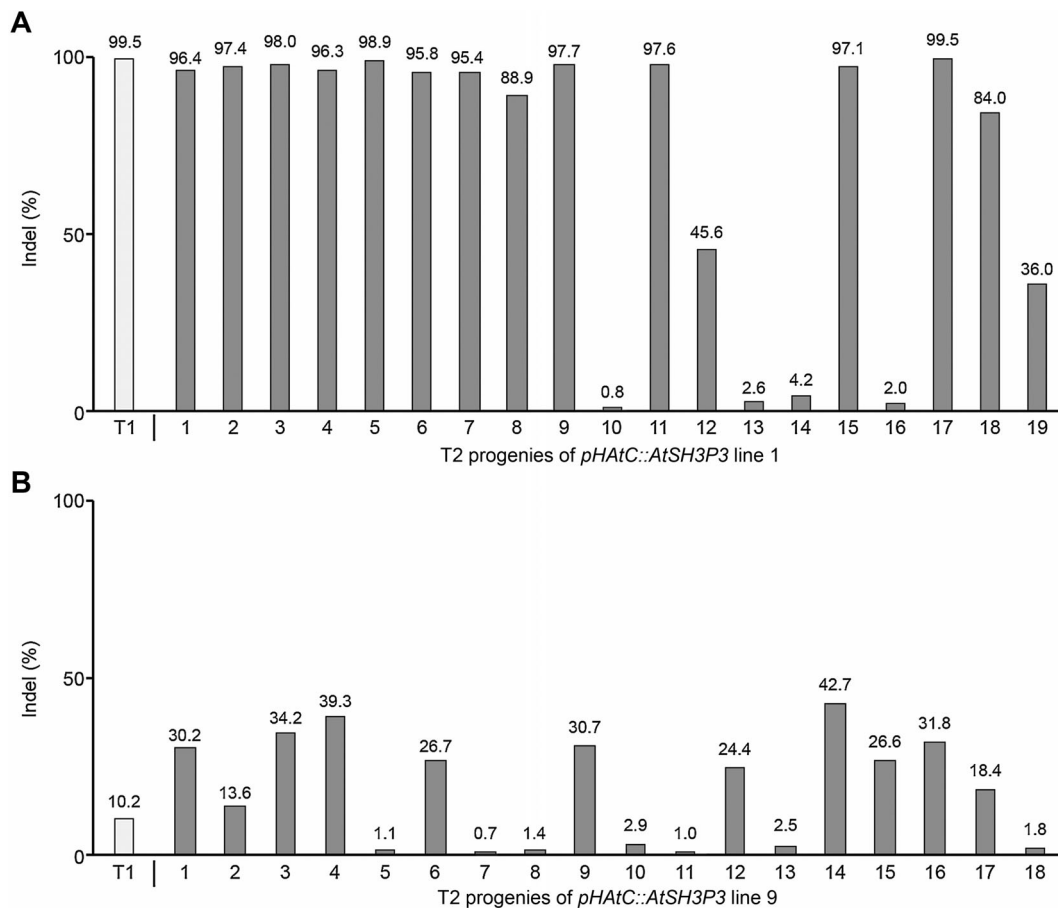


Figure 4. Heritable target gene editing in *Arabidopsis thaliana*

The mutation ratios (indel %) in T2 progenies of T1 *A. thaliana* transgenic lines (line 1 and line 9) harboring *pHAtC::AtSH3P3-sgRNA1*. Multiple leaves from each T2 plant were collected during the initiation of stem elongation, and targeted next-generation sequencing (NGS) analysis was performed to calculate the indel ratio of T2 progenies (See Materials and Methods and Figure 3).

The indel ratio in *A. thaliana* T1 plants was 99.5% and 10.2%, and those in *N. attenuata* T0 plants was 33.5% and 8.6%. These mutation percentages are similar with those reported in *Arabidopsis* T1 transformants (Mao et al. 2016) and barley T0 transformants (Lawrenson et al. 2015). The mutation ratios that we observed in somatic tissues of *A. thaliana* T1 plants were largely correlated with mutation ratios in the next generation (Figure 4). When the mutation ratio in T1 plant was close to 100%, most T2 plants originating from this T1 plant were close to 100%. However, some offspring have less than 5% mutation ratios, demonstrating that the high mutation ratio in T1 somatic tissues do not guarantee that in T2 generation. Major mutation patterns in *A. thaliana* and *N. attenuata* leaves were a single-base insertion at the cleavage site (Figures 3, S3), but we need more comprehensive analysis to conclude the overall indel pattern in edited plants employed by the CRISPR-Cas9 system.

Several binary vector systems for *Agrobacterium*-mediated plant transformation have been developed to express Cas9 protein and an sgRNA in a plant cell. While most of the systems use one or two CaMV 35S promoters to express transcripts coding for Cas9 protein (Belhaj et al. 2013), a developing embryo-specific promoter of *INCURVATA2* or a

gem line-specific promoter of *SPOROCTELESS* was fused to Cas9 coding sequence to enhance heritable targeted mutations in *A. thaliana* (Hyun et al. 2014; Mao et al. 2016). The location of NLS motifs in Cas9 protein varies among constructs; NLSs are located in both the N- and C-terminus or only C-terminus of Cas9 protein. However, there is little information on how the NLS location affects the localization of Cas9 protein and the efficiency of genome editing. There is no doubt that a single vector system expressing several sgRNAs is convenient to induce mutations in several target genes. For this purpose, Golden Gate ligation and Gibson Assembly have been used to assemble several U6::sgRNA cassettes and insert them into a single binary vector (Xing et al. 2014; Ma et al. 2015). Alternatively, tRNA-processing system was used to make several sgRNA constructs under a single U6 promoter (Xie et al. 2015). While we focused on the targeted mutagenesis using one sgRNA, the assembled U6::sgRNA or tRNA-sgRNA-tRNA cassette can be cloned into our vector system. The pH(B)AtC construct currently contains *Arabidopsis* U6-26 promoter sequence (Hyun et al. 2014), but it can be easily replaced with any plant-specific U6/U3 promoter, which facilitates targeted genome editing in non-model plants. In addition, this GatewayTM compatible

system can be easily adapted to various applications with Cas9 derivatives. This vector system is versatile and can be applied to model plants and major crops as well as non-model plants.

MATERIALS AND METHODS

Vector constructions

We used the Gateway™ destination vector, pB2GW7 and pH2GW7 binary vectors to construct pBATC and pHATC (Karimi et al. 2007). The nomenclature of the designed binary vectors was as follows: B, Basta™ (Phosphinothricin) resistance gene; H, Hygromycin resistance gene; At, *Arabidopsis thaliana* U6 promoter driven sgRNA cassette; C, Human codon-optimized SpCas9 gene containing a nuclear localization signal (NLS) and a HA epitope tag at the C terminus of SpCas9 (Cho et al. 2013; Hyun et al. 2014).

The AtU6-26 promoter was amplified by PCR using Phusion High-Fidelity DNA polymerase (Finnzymes, Thermo Scientific, Waltham, MA, USA) from pYB200 (Hyun et al. 2014) with the following primers: AtU6-26-F, 5'-GGCTGAGCTCGA-ATTCGAATGATTAGGCATCGAACC-3'; AtU6-26-R, 5'-GGCGAGCT-CAAAAAGCACCGACTCGGTG-3'. Each U6-sgRNA cassette was inserted in the SpeI site in pB(H) 2GW7. To express Cas9 protein in plants, the human codon-optimized Cas9 including an NLS from p35-Cas9hc plasmid (Cho et al. 2013) was cloned into the entry vector, pDOR221 by the BP Clonase™ II enzyme mix (Invitrogen, Carlsbad, CA, USA). Cas9 cassette was subsequently transferred into pB(H) 2GW7 by the LR Clonase™ II enzyme mix (Invitrogen). Two binding sites of AarI restriction enzyme (Thermo Fisher Science) were designed to insert the oligonucleotides of target binding sequence into U6::sgRNA cassette. To achieve the unique AarI site in an sgRNA cloning, we removed the other AarI binding sites in the plasmid backbone and the SpCas9 by integrating a point mutation. Unique EcoRI and XhoI restriction sites in the binary vector were designed for the easy exchange of U6 promoter of the target plant genome. All ligation reactions in this study were done at room temperature with T4 DNA ligase (New England Biolabs, Ipswich, MA, USA). The complete sequence of pB(H)AtC was validated by Sanger Sequencing. pBATC (KU213970) and pHATC (accession number, KU213971) vectors have been deposited in the GenBank database.

The following conditions were used for the annealing of the mixture containing 100 µmol/L of each oligonucleotide (Figure 2B): 95 °C for 5 min, linear gradient from 95 °C to 25 °C for 70 min, and 10 °C to maintain annealed product integrity.

Agrobacterium-mediated transformation

A. thaliana Col-o ecotype was used for *Agrobacterium tumefaciens* GV3101-mediated plant transformation via the floral dip (Clough and Bent 1998). Col-o plants were grown under long day conditions (16 h light/8 h dark) at 22 °C with ± 1 °C in a growth room (Koencon, Hanam, South Korea). Light was generated from a 32 W Osram lamp (170 µmol/2 m/s).

N. attenuata Utah ecotype seeds were provided by the Department of Molecular Ecology at the Max Planck Institute for Chemical Ecology, which were originally collected from a population in southwestern Utah, in the USA. Seeds were

sterilized and germinated in Gamborg B5 medium (Duchefa Biochemie, Haarlem, The Netherlands), as described in Krügel et al. (2002). *N. attenuata* plants were grown under long day conditions (16 h light/8 h dark) at 25 °C with ± 1 °C in a growth room (Koencon, Hanam, South Korea). We used hypocotyl tissues for *A. tumefaciens*-mediated transformation (Krügel et al. 2002).

Targeted deep sequencing and mutation pattern analysis

The genomic DNA for targeted deep sequencing analysis was extracted from randomly selected two or more leaves in each individual of T1 or T2 plants. The on-target sequence was amplified from genomic DNA using target-specific primers (Table S1) that we designed from reference genome sequences. Multiplexing indices and sequencing adaptors were added by additional PCR using the protocol supplied from the sequencing company, Macrogen (Seoul, South Korea). High-throughput sequencing was performed using Illumina Miseq (v2, 300-cycle; San Diego, USA) with the paired-end multiplexed library.

Raw reads of paired-end Miseq sequencing were joined by the program, 'fastq-join' implemented in the package, 'eu-util' (<http://code.google.com/p/ea-utils>). Mutation counts and patterns in joined reads were analyzed using a personal python script.

We also used several randomly selected leaves from T2 individuals and amplified the on-target region using the same primer sets used in the targeted deep sequencing. We directly sequenced the PCR product using Sanger sequencing method in the company Macrogen (Seoul, South Korea)/ Bioneer (Daejeon, South Korea).

ACKNOWLEDGEMENTS

We thank Emily Wheeler, Boston, for editorial assistance. This work was supported by Institute for Basic Science (IBS-R021-D1).

AUTHOR CONTRIBUTIONS

H.K., J.R., M.-K.C., J.K., S.-G.K., J.K. designed experiments and made the vectors; B.-C.K., H.-M.A., S.B. performed plant transformation and tissue culture; H.K. and S.-T.K. analyzed deep sequencing data; H.K., S.-T.K., S.-G.K. mainly wrote the manuscript; J.-S.K., S.-G.K., H.K. and S.-T.K. contributed to the revision of the manuscript and oversaw the project.

REFERENCES

- Araki M, Ishii T (2015) Towards social acceptance of plant breeding by genome editing. *Trends Plant Sci* 20: 145–149
- Bae S, Kweon J, Kim HS, Kim JS (2014) Microhomology-based choice of Cas9 nuclease target sites. *Nat Methods* 11: 705–706
- Baltes NJ, Voytas DF (2015) Enabling plant synthetic biology through genome engineering. *Trends Biotech* 33: 120–131
- Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V (2013) Plant genome editing made easy: Targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* 9: 39

- Bibikova M, Golic M, Golic KG, Carroll D (2002) Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. **Genetics** 161: 1169–1175
- Brooks C, Nekrasov V, Lippman ZB, Van Eck J (2014) Efficient gene editing in Tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. **Plant Physiol** 166: 1292–1297
- Cho SW, Kim S, Kim JM, Kim J-S (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. **Nat Biotech** 31: 230–232
- Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, Kim JS (2014) Analysis of off-target effects of CRISPR Cas-derived RNA-guided endonucleases and nickases. **Genome Res** 24: 132–141
- Clough SJ, Bent AF (1998) Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. **Plant J** 16: 735–743
- Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. **Nature** 471: 602–607
- Doudna JA, Charpentier E (2014) The new frontier of genome engineering with CRISPR-Cas9. **Science** 346: DOI: 10.1126/science.1258096
- Guilinger JP, Thompson DB, Liu DR (2014) Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. **Nat Biotech** 32: 577–582
- Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. **Cell** 157: 1262–1278
- Hyun Y, Kim J, Cho SW, Choi Y, Kim JS, Coupland G (2014) Site-directed mutagenesis in *Arabidopsis thaliana* using dividing tissue-targeted RGEN of the CRISPR/Cas system to generate heritable null alleles. **Planta** 241: 271–284
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. **Science** 337: 816–821
- Karimi M, Depicker A, Hilson P (2007) Recombinational cloning with plant gateway vectors. **Plant Physiol** 145: 1144–1154
- Karimi M, Inzé D, Depicker A (2002) GATEWAYTM vectors for *Agrobacterium*-mediated plant transformation. **Trends Plant Sci** 7: 193–195
- Kim D, Bae S, Park J, Kim E, Kim S, Yu HR, Hwang J, Kim JI, Kim JS (2015a) Digenome-seq: Genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. **Nat Methods** 12: 237–243
- Kim H, Kim JS (2014) A guide to genome engineering with programmable nucleases. **Nat Rev Genet** 15: 321–334
- Kim H, Kim ST, Kim JS, Kim SG (2015b) Targeted genome editing for crop improvement. **Plant Breed Biotech** 3: 283–290
- Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales APW, Li Z, Peterson RT, Yeh JRJ, Aryee MJ (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. **Nature** 523: 481–485
- Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, Joung JK (2016) High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. **Nature** 529: 490–495
- Krügel T, Lim M, Gase K, Halitschke R, Baldwin IT (2002) *Agrobacterium*-mediated transformation of *Nicotiana attenuata*, a model ecological expression system. **Chemoecology** 12: 177–183
- Lawrenson T, Shorinola O, Stacey N, Li C, Østergaard L, Patron N, Uauy C, Harwood W (2015) Induction of targeted, heritable mutations in barley and *Brassica oleracea* using RNA-guided Cas9 nuclease. **Genome Biol** 16: 258
- Li JF, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J (2013) Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. **Nat Biotech** 31: 688–691
- Ma X, Zhang Q, Zhu Q, Liu W, Chen Y, Qiu R, Wang B, Yang Z, Li H, Lin Y, Xie Y (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. **Mol Plant** 8: 1274–1284
- Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. **Nat Biotech** 31: 833–838
- Mao Y, Zhang Z, Feng Z, Wei P, Zhang H, Botella JR, Zhu JK (2016) Development of germ-line-specific CRISPR-Cas9 systems to improve the production of heritable gene modifications in *Arabidopsis*. **Plant Biotech J** 14: 519–532
- Mojica FJM, Diez-Villasenor C, Garcia-Martinez J, Almendros C (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. **Microbiology** 155: 733–740
- Nekrasov V, Staskawicz B, Weigel D, Jones JDG, Kamoun S (2013) Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. **Nat Biotech** 31: 691–693
- Park J, Bae S, Kim J (2015) Cas-Designer: A web-based tool for choice of CRISPR-Cas9 target sites. **Bioinformatics** 31: 4014–4016
- Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. **Cell** 152: 1173–1183
- Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV (2015a) *In vivo* genome editing using *Staphylococcus aureus* Cas9. **Nature** 520: 186–191
- Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F (2015b) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. **Cell** 154: 1380–1389
- Rouet P, Smih F, Jasin M (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. **Mol Cell Biol** 14: 8096–8106
- Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL, Gao C (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. **Nat Biotech** 31: 686–688
- Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan, WX, Zhang, F (2016) Rationally engineered Cas9 nucleases with improved specificity. **Science** 351: 84–88
- Sternberg SH, Doudna JA (2015) Expanding the biologist's toolkit with CRISPR-Cas9. **Mol Cell** 58: 568–574
- Sun X, Hu Z, Chen R, Jiang Q, Song G, Zhang H, Xi Y (2015) Targeted mutagenesis in soybean using the CRISPR-Cas9 system. **Sci Rep** 5: 10342
- Voytas DF, Gao C (2014) Precision genome engineering and agriculture: Opportunities and regulatory challenges. **PLoS Biol** 12: e1001877
- Wang S, Zhang S, Wang W, Xiong X, Meng F, Cui X (2015) Efficient targeted mutagenesis in potato by the CRISPR/Cas9 system. **Plant Cell Rep** 34: 1473–1476
- Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, Qiu JL (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. **Nat Biotech** 32: 947–951
- Xie K, Minkenberg B, Yang Y (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. **Proc Natl Acad Sci USA** 112: 3570–3575

Xing HL, Dong L, Wang ZP, Zhang HY, Han CY, Liu B, Wang XC, Chen QJ (2014) A CRISPR/Cas9 toolkit for multiplex genome editing in plants. **BMC Plant Biol** 14: 327

Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. **Cell** 163: 759–771

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Figure S1. Map of RGEN binary vector pHA_tC

Figure S2. Map of RGEN binary vector pBA_tC

Figure S3. Sanger sequencing results for T2 progenies from pHA_tC::AtSH3P3-sgRNA1 line 1

The targeted genomic region was amplified by high fidelity Phusion-PCR reaction and sequenced with Sanger sequencing method to examine the major mutation patterns in T2 somatic tissues. Base calls with the chromatography were kept as suggested by sequencing company and the alignment was carried out in the program, 'Geneious' (BIOMATTERS LIMITED, New Zealand).

Table S1. Primer list used for targeted deep sequencing